DOCUMENT RESUME

ED 151 985 EA 010 491

AUTHOR Shapson, Stan M.; And Others
TITLE Results of an Experimental Study of the Effects of Class Size.
SPONS AGENCY Ontario Dept. of Education, Toronto.
PUB DATE Mar 78
NOTE 39p.; Paper presented at the Annual Meeting of the American Educational Research Association (Toronto, Ontario, March 27-31); Appendix A may not reproduce clearly due to very small print of original
AVAILABLE FROM Ontario Ministry of Education, Ontario Government Bookstore, 880 Bay Street, Toronto, Ontario, Canada M5S 1Z8 ($2.00)

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS *Academic Achievement; *Class Size; Conventional Instruction; *Educational Research; Intermediate Grades; Self Concept; Statistical Data; *Student Attitudes; *Student Teacher Relationship; Tables (Data); *Teacher Attitudes; Teaching Styles

ABSTRACT

This study examines experimentally the differences among four class sizes ranging from 16 to 37 pupils. Sixty-two grades 4 and 5 classes in 11 schools in Toronto participated in the study. The study investigated the effects of class size on teachers' expectations about the effects of specific class sizes; the attitudes and opinions of students, teachers, and parents; student achievement in reading, mathematics, composition, and art; the academic self-concept of students; and a variety of classroom process variables. The data indicate that varying class size in these two grades resulted in few changes in classroom functioning or in pupil achievement, except in acquisition of mathematics concepts. It is easy to conclude that reducing class size is perceived as beneficial by teachers. Students' attitudes toward school, however, are unlikely to change; neither are teachers' instructional styles. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED151985

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

RESULTS OF AN EXPERIMENTAL STUDY OF THE EFFECTS OF CLASS SIZE[1]

Stan M. Shapson[2] Simon Fraser University

Edgar N. Wright)
Gary Eason) - Toronto Board of Education
John Fitzgerald)

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Stan Shapson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

A paper presented at the annual conference of the American Educational Research Association, March 27 - 31, 1978, Toronto, Canada

---

1. The research project, whose results are presented in this paper was funded under contract by the Ministry of Education, Ontario

2. Dr. Shapson was formerly Research Associate for the Toronto Board of Education.

EA 010 491

RESULTS OF AN EXPERIMENTAL STUDY OF THE EFFECTS OF CLASS SIZE

S.M. Shapson
E.N. Wright
G. Eason
J. Fitzgerald

Abstract

Most of the previous research on the effects of class size has been non-experimental; their results often have been difficult to interpret because important variables have not been controlled and there has been little consistency between these studies on the definitions of "large" and "small" classes. This study examined experimentally the differences between four class sizes which ranged from size 16 (the size of many special education classes) to size 37 (the approximate size of the largest regular classes in Toronto schools). Specifically this study, investigated the effects of class size on - teachers' expectations about the effects of specific class sizes; the attitudes and opinions of students, parents, and teachers; student achievement in reading, mathematics, composition, and art; the academic self-concept of students; and a variety of classroom process variables (e.g., teacher-pupil interaction, pupil participation).

Grade 4 and 5 classes at 11 schools in Metropolitan Toronto took part in the two year study. In the first year, teachers and Grade 4 students were randomly assigned to classes of size 16, 23, 30 or 37. The student assignments were stratified by sex and by academic performance. In the second year, the same teachers and students were similarly assigned to Grade 5 classes, with the constraints that students should not be in a class size of 16 or 37 for both years of the study, and that teachers who taught classes of the two larger sizes receive classes of the two smaller sizes. Class sizes were monitored so that they did not vary by more than ± 3 from the assigned class size.

Participating teachers had definite expectations about the effects of specific class sizes; they expected the smaller classes (i.e., 16 and 23) to have many advantages and the larger classes (30 and 37) to have many disadvantages. Teachers' experience during the study confirmed most of their initial expectations. For example, teachers expected that there would be more individualization in the smaller classes, and after teaching a small class reported that there had been more individualization.

However, most of the other data failed to support teachers' opinions. It was observed that teachers spent as much time talking to the entire class in the small classes as in large classes, and spent no more time talking to individual pupils in small classes. Pupils, however, were more likely to be addressed individually by the teacher in the smaller classes; this difference occured directly because the number of pupils in the class was reduced and the teacher continued to spend the same

amount of time talking individually to pupils. No other aspect of teacher-pupil interaction was affected by class size nor was pupils' participation. Reading was taught more frequently in classes of 16 than in classes of other sizes.

Students' attitudes and self-concepts were not affected by class size. However, scores on the mathematics-concepts subtest of the Canadian Tests of Basic Skills were significantly higher in classes of 16 than in classes of 30 or 37. There were no effects of class size for the other student achievement measures (reading, vocabulary, mathematics-problem solving, art and composition).

An important characteristic of this study which limits generalization of the results is that teachers were given no training in the use of appropriate methods to accomplish specified educational objectives with different class sizes. Essentially, this study changed class size and nothing else. The results also cannot be safely generalized beyond the junior grades nor to classes operating on team-teaching, or on rotary timetables.

RESULTS OF AN EXPERIMENTAL STUDY OF THE EFFECTS OF CLASS SIZE

S.M. Shapson
E.N. Wright
G. Eason
J. Fitzgerald

The issue of class size has long attracted the interest of the educational community. A consistent theme in discussions of class size has been teachers' expressed preference for small classes. However, research on the effects of class size has not produced such consistent results as have surveys of teachers' opinions and attitudes. Blake's (1954) extensive review of the class size literature revealed two striking problems: the difficulty of conducting scientifically adequate research on class size and the inconsistent findings between studies. For example, after locating 267 studies, Blake discovered that only 85 had used research procedures, and that only 22 satisfied his criteria of scientific adequacy. Of the latter 22 studies, the results were distributed among those favouring small classes, those favouring large classes with others showing no effects due to class size.

More recently there has been an attempt to pinpoint the major problems and issues concerning research on class size (e.g. Lindbloom, 1970; Ryan and Greenfield, 1975; Shapson, 1972). Briefly, some highlights of the previous research are as follows:

-- most class size research has been non-experimental; data were collected in existing situations with no attempt to experimentally set up different class sizes. Results of these studies have been inconsistent

and difficult to interpret because many important variables have not been controlled.

-- the available research also indicates that other educational variables (e.g., the teacher, and the method of instruction) are more important than class size.

-- in the few cases where attempts were made to experimentally manipulate the size of classes, the studies were not well-designed (e.g., the large class was given to the most experienced teacher or they included such a small number of classes that the results are readily attributable to teacher effects).

-- there has been little consistency between studies on the definitions of "large" and "small" classes. A class size (35 for example), that in one study is viewed as small (compared to size 50) turns out to be large in another study (compared to size 25).

In the early 1970's, diminishing school enrolments and provincial ceilings on educational expenditures were two of the factors that triggered a new round of debate on class size in Ontario. The Toronto Teachers' Federation passed a motion requesting that the Board of Education for the City of Toronto institute "a research project to determine the effects of class size on the education of a child in Toronto." Subsequently, the research project whose results are described in this paper was funded under contract by the Ontario Ministry of Education (Wright, Shapson, Eason and Fitzgerald, 1977).

The study was designed to examine experimentally the differences between four class sizes which ranged from size 16 (the size of many

special education classes) to size 37 (the approximate size of the largest regular classes in Toronto schools). Specifically the effects of class size were investigated on: teachers' expectations about the effects of specific class sizes; the attitudes and opinions of students, parents, and teachers; student achievement in reading, vocabulary, mathematics, composition, and art; the academic self-concept of students; and a variety of classroom process variables (e.g., teacher-pupil interaction, pupil participation, method of instruction, subject emphasis, use of educational resources, classroom atmosphere).

The purpose of this paper is to present the main findings of the study. Since the issues and details concerning the design and implementation of the study is the focus of a paper which was delivered earlier in this session (Wright et al 1978), only the main methodological points will be summarized in the following section.

METHODOLOGY

Sample:

Sixty-two Grade 4 and 5 classes in 11 schools in Metropolitan Toronto participated in the study.

In all schools efforts were made to ensure that teachers had the option of participating on the basis of informed consent. In addition, only teachers who had at least two years teaching experience were selected; the mean number of years taught was 7.2 with 50% of the participants having taught for five years or less. Approximately 70% of the sample were women as is the case with the system-wide figure for elementary school teachers.

Although students from all socio-economic levels were represented, there was a higher proportion of students in the sample from the lower socio-economic categories in comparison to the entire Toronto school population; 52.5% of the sample compared to 43.8% of the school population represent the lower socio-economic categories (Blishen, 1967--categories 25-31.99). Seventy-three percent (73%) of the students were born in Canada (vs. approximately 70% for the elementary school population in Toronto).

Instrumentation

The Toronto Classroom Observation Schedule (TCOS). The TCOS was designed, field tested and found to be adequately reliable prior to the start of the study. The schedule includes a fixed time-sampled observation system (investigating teacher verbal behavior, pupil participation and pupil aggressive behavior); an observation checklist (method of instruction,

subject emphasis, use of educational aids and physical conditions), and a Classroom Atmosphere Rating scale. The following variables were investigated with the TCOS:

(1) Teacher-Pupil interaction--measures three aspects of teachers' verbal behavior: content (whether verbalization deals with pupils' behavior, course content or routine classroom procedures); function (approving, disapproving or neutral verbalizations--verbalizations about course content are further categorized as questioning or telling); and audience (whether the verbalization is addressed to an individual, group or the whole class).

(2) Student Participation--measured in two ways, by observing both the whole class and randomly selected individual pupils. The participation of individual pupils is classified as verbal or non-verbal, with verbal participation further classified according to whether it is self-initiated or reactive, and by target audiences (i.e. teacher, peer or self). The pupil can also be classified as off task or as having no assigned task.

(3) Pupil Satisfaction--the frequency of all overt, disruptive, hostile acts occurring within a fixed time period; and a count of the number of different pupils acting aggressively.

(4) Method of Instruction--a ranking of the frequency of various methods of instruction.

(5) Subject Emphasis--a ranking of the subjects taught according to the amount of time spent on them.

(6) Physical Conditions--checklist of classroom noise, furniture arrangement and educational aids.

(7) Use of Educational Aids--the proportion of pupils using audio-visual, written and mechanical or concrete aids.

(8) Classroom Atmosphere--ratings on items of classroom functioning representative of classroom atmosphere (e.g., the pupils' regard for the teacher).

Indicators of Quality. This observation instrument, developed at Columbia University is intended to measure "quality of education," along four categories of classroom activity: individualization, inter-

personal regard, creative expression and group activity (see Olson & Vincent, 1972). It consists of 51 items, each of which describes one "positive" and one "negative" classroom activity (or "sign"). The most commonly used scores are the number of positive signs, the number of negative signs, and the difference between the numbers of positive and negative signs.

The Canadian Tests of Basic Skills (T. Nelson, Canada Ltd., 1968). Four subscales of this standardized achievement test were administered: Vocabulary (17 minutes); reading comprehension (55 minutes); mathematics concepts (30 minutes); and mathematics problem solving (30 minutes).

Art Sample. Students' art work on the topic "What I like to Do" was collected under standard procedures. The art samples were rated on a five-point scale which reflected developmental stages from the manipulative stage (code 1) through to the preadolescent stage (code 5). The rating categories were largely based on previous work by Pennington (1967) and Gaitskell & Harwitz, (1970).

Composition Sample. Compositions on the topics "Dreams" (the first year) and "Wishes" (the second year) were collected under standard procedures. A five point rating scale from a lower (code 1) "Illegible" category to an upper (Code 5) "Superior" category was developed for assessing the compositions based on previous work (Wright and Reich, 1972).

Self-Concept. The North York Self-Concept Inventory, developed by Shapson, Virgin and Crawford (1971), contains thirty items selected to measure the self-concept of pupils with respect to their school

environment. Students indicate whether they feel each statement is "true" or "not true" and a score of "one" is assigned to each response indicating a positive self-concept. The test/retest reliability coefficient was found to be .81 in previous studies.

Pupils' Attitudes Toward School Questionnaire. This questionnaire which was designed, pilot tested and revised for use in the study, contains 36 items selected to measure pupils': happiness and satisfaction; reaction to the physical and social classroom environment; attitudes towards teachers and other pupils; and attitudes toward subject taught.

Teachers' Questionnaires. A background information questionnaire which was administered before teachers were informed of their assigned class size, was used to describe the characteristics of the participating teachers (e.g., previous experience) and enable teachers to express their expectations of each of the proposed class sizes. A second questionnaire was designed to survey a wide range of teachers' attitudes and opinions toward their assigned class size.

Semantic Differential. This technique of assessing attitudes involves rating a concept, such as "My Classroom" according to pairs of polar objectives (e.g., good... bad) which are separated by five or seven point scales. Two forms of a seven point scale semantic differential were developed for participating teachers to describe "My Classroon," and "The Pupils I Teach," respectively. A five point semantic differential was used with pupils to describe "My Classroom."

Parent Questionnaire. A questionnaire was designed to survey parental opinions about class size and its possible effects on their children. (These data are not included in this paper).

Overview of Design and Procedure

In the first year of the study, teachers and Grade 4 students were randomly assigned to classes of four sizes: 16, 23, 30 or 37. The student assignments were stratified by sex and by academic performance. At the beginning of the second year the same teachers and students were similarly assigned to Grade 5 classes, with the constraints that students should not be in a class size of 16 or 37 for both years of the study, and that teachers who taught classes of the two larger sizes receive classes of the two smaller sizes (and vice-versa).

The actual size of each participating class was closely monitored during the year to ensure that the enrolment did not vary by more than ±3 from the assigned class size.

Paper and pencil measures of opinions and attitudes of participants (students, teachers and parents) were collected. Standardized achievement tests, a self-concept scale and an art and composition measure were administered to students each year. In addition, observations of a variety of classroom process variables were made with the Toronto Classroom Observation Schedule, which was designed for the study. This observation schedule was used during eight half-day visits to each class each year. "Indicators of Quality", a rating schedule used in previous research at Columbia Teacher's College (Olson, 1970) also was used by observers for five twenty-minute visits to each class. A schedule of the data collection activities is attached in Appendix A.

Data Analysis

The final analysis included data from 62 classes--sixteen of each of class size 16 and 23, and fifteen of each of class size 30 and 37. In general, differences between class sizes were assessed by a one-way analysis of variance. Where a significant class size effect was found, pair-wise contrasts or range tests were conducted. In all cases, the unit of analysis was the class; that is, mean scores from each class were used rather than individual scores from each student or each observation visit.

For the observational data, the following procedure was followed:

-- means of each variable were compared (before the analysis of variance) with proportions tests to see if there were differences between years. Few differences were observed and all but one were extremely small.*

-- one-way analysis of variance of scores on each measure by class size were conducted.

-- when a significant difference was found, pseudovalues were calculated by Mosteller and Tukey' (1968) jackknife procedure and a similar analysis of variance was perofrmed on the psuedovalues.**

For the student data (affective scales and achievement measures) the following procedures were followed:

-- a one-way analysis of variance by class size was conducted.

-- the variability due to year of the study (i.e. grade 4 and grade 5) and teacher was removed using a multiple linear regression technique.

-- the analysis of variance was repeated on the "residual" data.

---

* Scores on the variable for which there was a large difference between years were adjusted (see the following presentation of results).

** The batching of data involved in this procedure helps to balance the effects of schools with different numbers of classes.

RESULTS

Teachers' Affective Measures

Teachers' Expectations. Prior to being assigned their class size, teachers were asked to indicate the effects they expected in each of the proposed class sizes. The responses were classified according to tone (positive, negative) and content (e.g., effects on student outcomes, environmental and teacher effects, program effects).

Teachers had positive expectations regarding effects of the smaller class sizes (16, 23) and negative expectations regarding the larger classes (30, 37). They felt that the smaller classes would enable them to give more attention to individual problems and to develop better rapport with pupils. On the other hand, they felt that the larger classes would result in less individual attention. Many teachers suggested that the smaller classes would allow for an individualized program (vs. one aimed at the "average" pupil in the larger classes), more remedial help and enrichment, more extracurricular materials and field trips.

Another noteworthy theme related to student outcomes; teachers expected academic improvement and the development of pupil self-confidence and independence in the smaller classes (vs. little involvement, and dependence in the larger classes). Teachers were also looking forward to a more relaxed, enjoyable environment in the smaller classes, in contrast to more structure and formality in the larger classes. A summary of teachers' expectations prior to the study is presented in Appendix B.

In order to determine if direct experience with a class size altered their expectations, teachers were again asked for their expectations

toward the end of both the first and second years of the study. A teacher's initial expectations were compared to those expressed by the same teacher after experiencing a particular size and a list of expectations which were restated was compiled.

It was discovered that teachers' expectations remained quite stable. For example, 43 initial statements expressed the opinion that class size would specifically determine the amount of individualization possible. After teacher experience, 35 statements restated that either the smaller class sizes (i.e. 16 and 23) actually resulted in more individualized programs and more individual attention or that individualization was difficult in size 30 and almost impossible in size 37. Seven of eight teachers repeated their prediction that pupils' achievement would be higher in the small classes. Nevertheless, some expectations turned out to be less resistant to change. While fourteen teachers thought opportunities for enrichment, field trips and discussions would be more easily instituted in class sizes 16 and 23, only four subsequent affirmations were recorded.

Teachers' Opinions. A series of questions asked for teachers' opinions about their class size. The distribution of responses signified that, the smaller the class size, the more likely were teachers to feel that:

-- the quantity of available teaching aids was superior; χ^2 (6) = 17.4, $p < .01$;

-- the display area and blackboard area in the classroom were superior; χ^2 (6) >18 in both cases, $p < .01$

-- the floor space in the classroom was superior; χ^2 (6) = 38.4, $p < .001$.

-- fewer students had difficulty seeing or hearing;
χ^2 (3) 20 in both cases, $p < .001$.

Teachers also were asked whether they made any changes directly related to the size of their class in a variety of areas such as class management and supervision, for example. The largest percentage of teachers to indicate that they had made changes were those who experienced size 16 (77 percent), followed by size 37 (65 percent) with the fewest changes being reported by teachers in size 30 (37 percent). An examination of the data by area of change revealed that teachers reported making the most change in physical layout (82 percent), followed by evaluation (72 percent), class management (71 percent) and supervision (66 percent). Comments made by teachers regarding these areas are summarized below:

> Physical Layout. Teachers of size 16 commented that they were pleased with the amount of room available, especially for centres and activities and they reported that pupils had a choice of where to work and were free to move around; teachers of size 30 reported that grouping was beginning to get tight; and in size 37, it was reported that the rooms were crowded and desks took up much room.
>
> Evaluation. Teachers of sizes 16 and 23 were happy because marking took little time, and corrections were immediate. Some alternative methods were mentioned such as interviews and pupil self-evaluation and it was felt that evaluation could be more personal and extensive. In size 30, it was reported that marking becomes more formal, time-consuming and is sometimes delayed. In size 37, marking and evaluation was reported to be overwhelming. Teachers enlisted pupils for marking. Evaluation was based on tests or specific projects, rather than on general teaching knowledge of pupils. Teachers seem dissatisfied with this kind of evaluation.
>
> Management. Teachers of size 16 were happier because they were able to run their classrooms more smoothly and efficiently. Children were more responsible in all areas, including working on their own, handling supplies and behaving appropriately. In size 37, it was reported

that rules had to be enforced. Management was more formal with restricted physical movement of students.

Supervision. In size 16 it was reported that there was little need for the teacher to supervise; often the children made supervisory decisions. The feeling of being more tolerant of children's actions, was evident in the majority of replies. In size 37, it was suggested that teachers had to police behavior. The classes were more structured, the teachers strict. There was no feeling of freedom for the teacher.

Teachers' opinions to several questions were analyzed as a function of two "size change" groups: one group of teachers (n = 15) who went from a large class size (30, 37) the first year to a small size (16, 23) the second year; and a second group of teachers (n = 13) who went from a small size to a large size. Teachers who went from a large to a small class size were significantly more likely to:

-- like their current class, χ^2 (1) = 35.2, $p < .001$;

-- report a higher personal energy level, χ^2 (2) = 76.4, $p < .001$;

-- feel that their pupils contributed more, paid more attention and were more satisfied this year; χ^2 (2)> 90 in each case, $p < .001$.

Attitude Scale. Teachers rated the concepts "My Classroom" and "The Pupils I Teach" on a seven point semantic differential scale (where seven was the positive end of the scale). The mean positive scale scores are presented by class size in Table 1.

A significant effect due to class size was found for the concept "My Classroom," $F(3,58) = 4.2$, $p < .01$. Pair-wise contrasts revealed that teachers who had experienced class size 16 rated "My Classroom" significantly more positively than teachers in size 30 ($t_{54} = 2.88$, $p < .01$) and than those in size 37 ($t_{54} = 3.10$, $p < .01$).

TABLE 1

MEAN SCORES ON TEACHERS' SEMANTIC DIFFERENTIAL

	Class Size				Anova	
Concept	16	23	30	37	df	F
My Classroom	100.9	96.8	92.2	91.4	3,58	4.2*
Pupils I Teach	114.1	108.9	110.8	110.2	3,55	0.5

* $p < .01$

Observation of Classroom Process Variables

Data for the variables on the Toronto Classroom Observation Schedule (TCOS) for which significant effects due to class sizes were detected are presented in Table 2.

Teacher-Pupil Interaction. The following variable concerned with the categorization of Teacher Verbal Behavior were analyzed (maximum mean score is 32).

1. the frequency of utterances about classroom routines;
2. the frequency of lecturing ("tell") to all audiences (individual, group or class);
3. the frequency of lecturing to the class;
4. the frequency of questioning;
5. the frequency of approving utterances about course content;
6. the frequency of disapproving utterances about course content;
7. the frequency of all utterances about course content;
8. the frequency of approving utterances about pupils' behavior;
9. the frequency of disapproving utterances about pupils' behavior;
10. the frequency of neutral utterances about pupils' behavior;
11. the frequency of all utterances about pupils' behavior.

TABLE 2

OBSERVATIONAL VARIABLES PRODUCING SIGNIFICANT CLASS SIZE EFFECTS

Area of Effect	Variable	Mean Score Class Size 16	23	30	37	F*	Jackknife F**
Teacher-pupil interaction	Proportion of pupils addressed as individuals	.43	.34	.30	.25	14.43	12.96
Method of instruction	Lecture*** (maximum = 8)	4.19	4.69	3.13	3.80	3.63	4.06
Method of instruction	Supervision by teacher while pupils are working*** (maximum = 8)	1.94	2.44	3.60	1.92	2.94	3.38
Subject emphasis	Reading	4.38	3.69	2.60	3.13	2.95	2.82
Use of educational resources	Number of written aids used*** (maximum = 40)	15.06	17.81	17.93	15.27	4.71	4.53
Use of educational resources	Proportion of available written aids used***	.40	.48	.48	.41	4.61	4.50

* At the .05 level the critical F ratio is 2.77

** At the .05 level the critical F ratio is 2.90

*** The reliability of this variable was not demonstrated (see Wright, Shapson, Eason & Fitzgerald, 1977 for criteria and data on reliability).

12. the frequency of all approving utterances;
13. the frequency of all disapproving utterances;
14. the frequency of all neutral utterances
15. the frequency of all observations about social matters;
16. the frequency of all utterances addressed to individual pupils;
17. the frequency of all utterances addressed to groups; (not to formal learning groups, such as in reading)
18. the frequency of all utterances addressed to the class;
19. the frequency of pupil talking (that is, talking which is part of learning);
20. the frequency of observations during which neither the teacher nor any pupils were talking;
21. the frequency of all observations during which the teacher was not talking (that is, the sum of variables 19 and 20.)

None of these variables was affected by class size. Of particular interest is the failure of class size to affect the lecturing and audience variables, the frequency of utterances about course content or the frequency of utterances about routines (variables 1, 2, 3, 7, 16, 17 and 18).

Further data on teacher-pupil interaction were collected during the observation of the participation of individual pupils. The measure obtained was the proportion of pupils observed whom the teacher addressed individually. There was a significant class size effect for this variable; pupils were more likely to be addressed individually in class size 16 than in the other three sizes, while pupils in class size 23 were more likely to be addressed individually than those in sizes 30 and 37. The linear trend accounted for 96 percent of the variance.

Pupil Participation. Data on pupil participation were collected in the sections of the TCOS which categorize Individual Pupil Participation (IPP) and general Pupil Participation (PP). The general PP section

provided an estimate of the proportion of students in the classroom who were participating while the following IPP variables were assessed (maximum mean scores are 16):

(1) the frequency of observations during which the pupil observed was participating (that is, performing the task assigned to him);

(2) the frequency of observations during which the pupil was participating non-verbally;

(3) the frequency of observations during which the pupil was participating verbally;

(4) the frequency of observations during which the pupil was not participating;

(5) the frequency of observations during which the pupil had no task to perform (this variable includes lining up to have work checked).

None of the above variables were affected by class size.

Pupil Satisfaction. It was originally intended to measure this area of effect by categorizing Pupil Aggressive Behaviour (PAB). However, overt pupil hostility was reported during only 66 of 1,390 PAB observations. A related variable on the TCOS was the number of observations during which no conflict between pupils was reported. There was no significant class size effect on this variable.

Method of Instruction.* On a checklist, the observer recorded the three most frequent teacher's activities during the observation visits. Counts were made of the number of times the following activities were mentioned in the course of a year: lecturing, organization of learning activities,

* During the development of the TCOS, the concept of method of instruction was broadened to include non-academic activities as well, so that "teacher activity" would probably be a better name for the area of effect investigated.

question and answer, and supervision. No significant differences between class sizes were found for organization, question and answer, and supervision. However, a significant difference was found for the lecturing variable; lecturing was recorded less frequently in class size 30 than in classes of other sizes. (The linear trend was not significant).

The teacher's use of time while pupils worked on their own was observed. The teacher was classified as working alone, supervising, working with an individual pupil, or working with a group of pupils. The number of times during a year that each of these categories was reported in each class was calculated (maximum score on each variable is eight). Only the supervision variable produced a significant class size effect; supervision by the teacher was found to be more frequent in class size 30 than in class size 16.

Subject Emphasis. The subjects studied during each observation visit were recorded and three categories--reading, mathematics and all language arts subjects (including reading)--were analyzed. The reading and mathematics variables were counts of the number of times during the year each subject was reported in each class (the maximum score was eight). The remaining variable was the mean number of language arts subjects reported during an observation (the maximum score was fifteen). The results revealed no significant class size effect for mathematics and language arts respectively. Reading was affected by class size; it occurred more frequently in class size 16 than in class size 30. However, the analysis of pseudovalues calculated by Mosteller and Tukey's jackknife procedure only approached significance; the probability of the F

value was .054.*

Educational Resources. The variables investigated were the number of audio-visual, written, and mechanical-concrete resources present in the classroom as well as those actually used. There was no significant class size effect for the presence of any of the above educational resources or for the use of audio-visual or mechanical-concrete aids. However there was a significant class size effect for the use of written aids; written aids were used more frequently in class sizes 23 and 30 than in size 16.

Classroom Atmosphere. There was no significant effect of class size on classroom atmosphere.

Indicators of Quality. Three scores were investigated--the number of positive signs recorded (the positive score), the number of negative signs recorded (the negative score) and the difference score. There was no significant effect of class size on Indicator of Quality Scores.

Student Affective Measures

Opinions of Subject Emphasis. Students were asked whether the amount of class time spent on core subjects (reading, writing and mathematics) and

---

* It was originally intended to measure the amount of time spent on each subject. However, a satisfactory measure could not be devised. It is difficult, for example, to decide how much time is spent on each subject when the teacher puts on the board a list of work to be done and the pupils spend the rest of the day working on subjects in which their work is not caught up.

non-core subjects (music, art, gym) was "too much", "okay" or "too little." For each subject, a mean score was calculated and a one-way analysis of variance (by class size group) was performed. Only for Art was a significant effect due to class size found, F (3, 1467) = 2.88, $p < .05$; in comparison to class size 16, students were significantly more likely to feel that too little time is spent on Art in class size 30 ($t = -2.77$, $p < .005$).

Attitudes Toward School Scale. This scale contained 30 items (18 positively worded, 12 negatively worded) to which students could "agree," "disagree" or be "undecided." A total scale score was obtained by scoring all items in the positive direction.

The mean attitude scale scores (unadjusted and residuals) are presented by class size in Table 3. First the analysis of variance of unadjusted mean scores revealed a significant effect apparently due to class size, F (3,58) = 3.65, $p < .05$; pair-wise contrasts revealed that attitudes toward school were significantly more positive for students in size 16 than for their peers in the other three class sizes respectively; $t(58) > 2.5$ in each case, $p < .001$.

However, when this analysis was repeated on the residual data after removing the variability due to year and teacher, there no longer was a significant effect due to class size, F (3,58) = 1.08, $p < .05$. Thus, the apparent class size effect noted initially was an artifact.

Semantic Differential Scale. Student attitudes toward the concept "My Classroom" were assessed through a semantic differential. The mean scale scores are also presented by class size in Table 3.

First an analyses of variance of unadjusted scores revealed no significant class size effect, $F(3,58) = 0.46$. To parallel the procedure used for the previous attitude scale, the analysis was repeated using the residual scores; again no difference was found between class sizes, $F(3,58) = 0.72$.

Self Concept. Also presented in Table 3 are the mean self-concept scores by class size. There was no significant effect of class size for student self-concept, $F(3,58) = 1.38$, $p < .05$.

TABLE 3

RESULTS OF PUPILS' AFFECTIVE MEASURES BY CLASS SIZE

Measure	Class Size				Anova	
	16	23	30	37	df	F
Attitudes Toward School Scale						
Unadjusted Mean Score	19.71	17.35	17.57	17.48	3,58	3.65*
Residual Mean Score	0.29	0.11	0.11	-0.08	3,58	1.08
Semantic Differential Scale						
Unadjusted Mean Score	8.63	8.93	8.78	9.20	3,58	0.46
Residual Mean Score	0.03	0.18	0.26	0.06	3,58	0.73
Self-Concept						
Unadjusted Mean Score	20.16	20.34	20.21	20.25		
Residual Mean Score	0.09	-0.08	0.13	-0.14	3,58	1.38

* $p < .05$

Student Achievement

The effects of class size were investigated on six achievement measures: art, composition, reading, vocabulary, mathematics-concepts and mathematics-problem solving. The mean scores (unadjusted values and residuals) are presented by class size in Table 4 for each measure.

The results reveal no significant differences attributable to class size for: art (F 3,58 = 1.29); composition (F 3,57 = 0.88); vocabulary (F 3,58 = 0.71); reading (F 3,58 = 0.16); and mathematics-problem solving (F 3,58 = 2.58); $p < .05$ in each case.

Only for mathematics-concepts, was there an overall significant effect due to class size (F 3,58 = 4.11), $p < .01$. Pair-wise contrasts revealed that students in class size 16 scored significantly higher in mathematics-computation than their peers in class size 30 ($t_{58} = 3.16$, $p < .005$) and in class size 37 ($t_{58} = 2.87$, $p < .01$). None of the other pair-wise contrasts were statistically significant.

TABLE 4

RESULTS OF STUDENT ACHIEVEMENT MEASURES BY CLASS SIZE

Measure	Class Size 16	23	30	37	Anova* df	F
Art						
unadjusted mean score	6.47	6.30	6.49	6.58		
(standard deviation)	(0.39)	(0.91)	(0.58)	(0.50)		
residual mean score	-0.07	0.01	-0.11	0.18	3,58	1.29
(standard deviation)	(0.24)	(0.59)	(0.35)	(0.49)		
Composition						
unadjusted mean score	4.56	4.40	4.51	4.46		
(standard deviation)	(0.77)	(0.85)	(0.66)	(1.01)		
residual mean score	0.14	0.02	-0.06	-0.11	3,57	0.88
(standard deviation)	(0.46)	(0.43)	(0.43)	(0.51)		
CTBS Vocabulary						
unadjusted mean score	43.16	43.58	42.42	42.84		
(standard deviation)	(9.55)	(9.97)	(6.51)	(8.06)		
residual mean score	-0.00	0.04	0.03	-0.07	3,58	0.71
(standard deviation	(0.22)	(0.24)	(0.24)	(0.14)		
CTBS Reading						
unadjusted mean score	44.08	44.47	44.05	44.28		
(standard deviation)	(7.99)	(8.62)	(6.00)	(7.22)		
residual mean	-0.00	-0.04	0.02	0.02	3,58	0.16
(standard deviation)	(0.30)	(0.30)	(0.27)	(0.25)		
CTBS Math-Concept						
unadjusted mean score	47.41	45.60	44.03	44.69		
(standard deviation)	(9.24)	(8.52)	(6.27)	(7.36)		
residual mean	0.15	0.00	-0.09	-0.07	3,58	4.11**
(standard deviation)	(0.22)	(0.20)	(0.18)	(0.23)		
CTBS Math-Problem Solving						
unadjusted mean score	51.17	49.33	48.20	48.82		
(standard deviation)	(8.83)	(8.21)	(6.36)	(6.84)		
residual mean score	0.14	0.04	-0.09	-0.10	3,58	2.57
(standard deviation)	(0.27)	(0.28)	(0.27)	(0.30)		

* Only data for the analyses of variance on the residual means are presented here. Similar analysis were concluded on the unadjusted means but none of these were significant at the .05 level.

** $p < .01$

SUMMARY AND DISCUSSION

This study collected substantial evidence which showed that varying class size in Grades 4 and 5 resulted in few changes in classroom functioning.

Of all the different variables examined in this study, the ones that were most affected by class size were the teachers' expectations before the study and their opinions and attitudes after experiencing another class size. The teachers expected the two smaller classes (16 and 23) to be different from the larger classes (30 and 37), especially in the amount of individualization possible. Their experiences with the smaller classes as they perceived it supported this expectation. Even though the observational data do not seem to give these findings much support, they must not be ignored; teachers do prefer to work in a smaller class.

The investigation of observational variables assessing teacher-pupil interaction showed very few effects of class size. Class size did not affect the amount of time teachers spent talking about course content or classroom routines; it did not affect their choice of audience for what they said--that is, teachers, when they changed sizes, did not alter the proportion of their time spent working with the class, with groups or with individuals. The differences in the frequency with which individual pupils were addressed by teachers were large, but as noted, there were no corresponding differences in the amount of time spent by the teacher talking to individual pupils. It seems, then, that pupils in classes of 16 and 23 had more individual interactions with their teachers simply because the amount of time spent talking to pupils individually was being distributed among fewer pupils. These findings are changes which generally would be

expected if class sizes were reduced and teachers did not change their style of teaching.

In addition, changing class size did not produce changes in pupils' participation. As well, the only observed difference in the use of educational resources was in the proportion of available written aids used. There were no significant differences in the quantities of educational resources seen in the different sizes. More interesting is the failure to confirm the teachers' opinions about the quantity of aids available in the different class sizes. Teachers tended to feel that the quantity of aids available in the smaller classes was superior, but the observational data report no differences in the numbers of aids available. The contradiction here may be only apparent, as the observational system counted only those aids which were displayed, rather than all aids which might have been available to the class. Another explanation is that the same number of resources may seem to be superior when there are fewer students to use them.

The only difference in method of instruction was found in the number of observation visits a year during which lecturing was used as a method of instruction. This particular variable, however, was not demonstrably reliable. A different assessment of the lecturing variable mentioned above, using a measure with demonstrated reliability, did not show any significant differences related to class size.

Reading was reported as having been taught during more observation visits in classes of size 16 than in classes of other sizes. However, the actual length of time spent on reading, and other subjects, could not be assessed, and students' subsequent achievement in reading was not affected

by class size.

Considering the preceding findings, it appears that, in general, teachers experience a big difference between the smaller, 16 or 23, and the larger, 30 or 37, class sizes. They report making many changes in physical layout, evaluation, class management and supervision as a result of the differences (they were not instructed or advised as to how they should adjust for the different sizes). Undoubtedly, to an individual, a change personally initiated, for whatever reason, is seen to be important and significant. Viewed externally these individual changes are either so small as to be undetectable, or else they are observable but still are relatively small when seen in the context of the wide variety of other teachers' actions and activities.

Student affective measures were also collected. Although initially it seemed that there were some class size effects, the apparent differences in student attitudes disappeared after introducing statistical controls to remove the variability due to teacher and year. Student opinions on subject emphasis were also collected and an effect related to class size was found for Art. Students in the small classes were less likely to say that Art received "too little" time than those in the larger classes. No significant class size effects could be identified in the students' measure of self-concept.

Standardized tests that were used to measure academic achievement showed a statistically significant class size effect only for Mathematics Concepts. There were no signficant differences for the measures of Reading, Vocabulary or Mathematics-Problem Solving. Students' development in Art and Composition also showed no differences between class sizes. An argument

against standardized tests was that performance in a creative endeavour such as Art or Composition would be more sensitive to class size effects than performance on a multiple choice test. This was not supported. These findings raise at least one possibility, namely that Mathematics is more sensitive to teacher availability than some of the other subjects.

Considering the data carefully, it is easy to conclude that reducing class size is perceived as beneficial by teachers. There will be a few other effects that will likely make the classroom a slightly better place for the students, whose Mathematics Concept skills might show an improvement, depending partly on how small the class becomes. Students' attitudes toward school, however, are unlikely to change, and the teachers' instructional style will not likely change very much either.

This study has limited itself to the effects of class size which are measurable within the framework of a traditional classroom organization, i.e., the effects of varying the number of pupils taught in one classroom by one teacher. It is essential to note that these findings cannot be generalized to all levels and kinds of education. This study was based on students in junior grades and the results cannot be applied to primary grades, secondary school or university. In addition at the higher levels of education, classes usually operate on rotary timetables and this study cannot be generalized to such a setting. The study also avoided mixing the concept of class size with open plan schools or team teaching, in which the traditional meaning of class size does not necessarily apply.

Obviously, the class size question has important financial considerations: decreasing class sizes, with all other factors held constant is an expensive proposition. An examination of the financial implications of

changing class size was beyond the scope of this study as was a cost analysis to compare the relative effectiveness of alternative components of school expenditures. There was no attempt in this study to examine different options for the same expenditure.

This study approached class size from the perspective of an analytical research question. Ryan and Greenfeld (1976) discuss different kinds of perspectives to the class size issue.

An important characteristic of this study which limits generalization of the results is that teachers were given no training in the use of appropriate methods to accomplish specified educational objectives with different class sizes. They were not given any particular instruction or advice regarding how they should adjust to the different sizes. However, it must be noted that if class size were to be changed by legislation, it is extremely unlikely that a well-defined in-service training program would be included in that legislation.

Essentially this study changed class size and nothing else. The researchers assumed that the influences of teachers and students were more important than class size and designed the study accordingly. Before further research is done it would be helpful to have clear definitions of educational objectives differentiated by class size and a set of strategies for attempting to reach them. As far back as 1932, Carr stated that before further work on class size be undertaken, the purpose of education should be clarified in order to give direction to subsequent investigation.

The authors expect to use the data collected in this study to go beyond the class size question and explore other relationships such as

those among teaching style, student attitudes and student performance. The next paper in this session is one of these endeavours (Fitzgerald, Wright, Eason and Shapson, 1978).

References

Blake, H.E. Class size: a summary of selected studies in elementary and secondary public schools. Unpublished doctoral dissertation, Teachers College, Columbia University, 1954.

Blishen, B.R. A socio-economic index for occupations in Canada. Canada Review of Sociology and Anthropology, 1967, 4(1), pp. 41-53.

Carr, W.G. New angle of attack needed on class size research. Nations Schools, 1932, 10, pp. 27-30.

Fitzgerald, J., Wright, E.N., Eason, G. & Shapson, S.M. Some secondary analyses of an experimental study of the effects of class size. Paper presented at the annual conference of the American Educational Research Association. Toronto, Canada, March 27-31, 1978.

Gaitskell, C., & Hurwitz, A. Children and their art: Methods for the elementary school. (2nd ed.) New York: Harcourt, Brace & World, 1970.

Lindbloom, D.H. Class size as it affects instructional procedures and educational outcomes. Minneapolis: Educational Research and Development Council of the Twin Cities Metropolitan Area, 1970. (ERIC Document Reproduction Service No. ED 059 532.)

Mosteller, F., & Tukey, J. Data analysis, including statistics. In G. Lindzey & E. Aronson (Eds.) The handbook of social psychology, (Second Edition, Volume 2), Reading, Mass.: Addison-Wesley, 1968, pp. 80-203.

Olson, M. Classroom variables that predict school system quality. IAR Research Bulletin, 1970, 11(1), pp. 1-11.

Pennington, S. Children's drawings: a selected review of literature. Toronto: The Board of Education for the City of Toronto, Research Department, 1967 (#43).

Ryan, D.W., & Greenfield, T.B. The class size question: development of research studies related to the effects of class size, pupil-adult, and pupil-teacher ratios. Toronto: Ministry of Education, 1975.

Ryan, D.W., & Greenfield, T.B. Clarifying the class size question: evaluation and synthesis of studies related to the effects of class size, pupil-adult, and pupil-teacher ratios. Toronto: Ministry of Education, 1976.

Shapson, S. Optimum class size? A review of the literature. Toronto: The Board of Education for the City of Toronto, Research Department, 1972 (#114). (Reprinted in Ontario Education, 1973, 5, pp. 18-24.)

Shapson, S.M., Virgin, A.E., & Crawford, P. Development of an instrument to measure self concept in schools. Toronto: Board of Education for the Borough of North York, Educational Research Services, October, 1971.

Vincent, W., & Olson, M. Measurement of school quality and its determiners. New York: Teachers College, Columbia University, 1972.

Wright, E.N., & Reich, C. Language: a study of fundamental skills. Toronto: Board of Education for the City of Toronto, Research Department, 1972, (#108).

Wright, E.N., Shapson, S.M. Eason, G. & Figzgerald, J. Effects of class size in the junior grades: a study. Toronto: Ministry of Education, 1977.

Wright, E.N., Shapson, S.M., Eason, G. & Fitzgerald, J. Issues concerning the design and implementation of an experimental study of the effects of class size. Paper presented at the annual conference of the American Educational Research Association. Toronto, Canada, March 3. 7-31, 1978.

32.

APPENDIX A

NUMBER AND CONTENTS OF THE TEACHERS' COMMENTS REGARDING EXPECTATIONS OF THE EFFECTS OF CLASS SIZE*

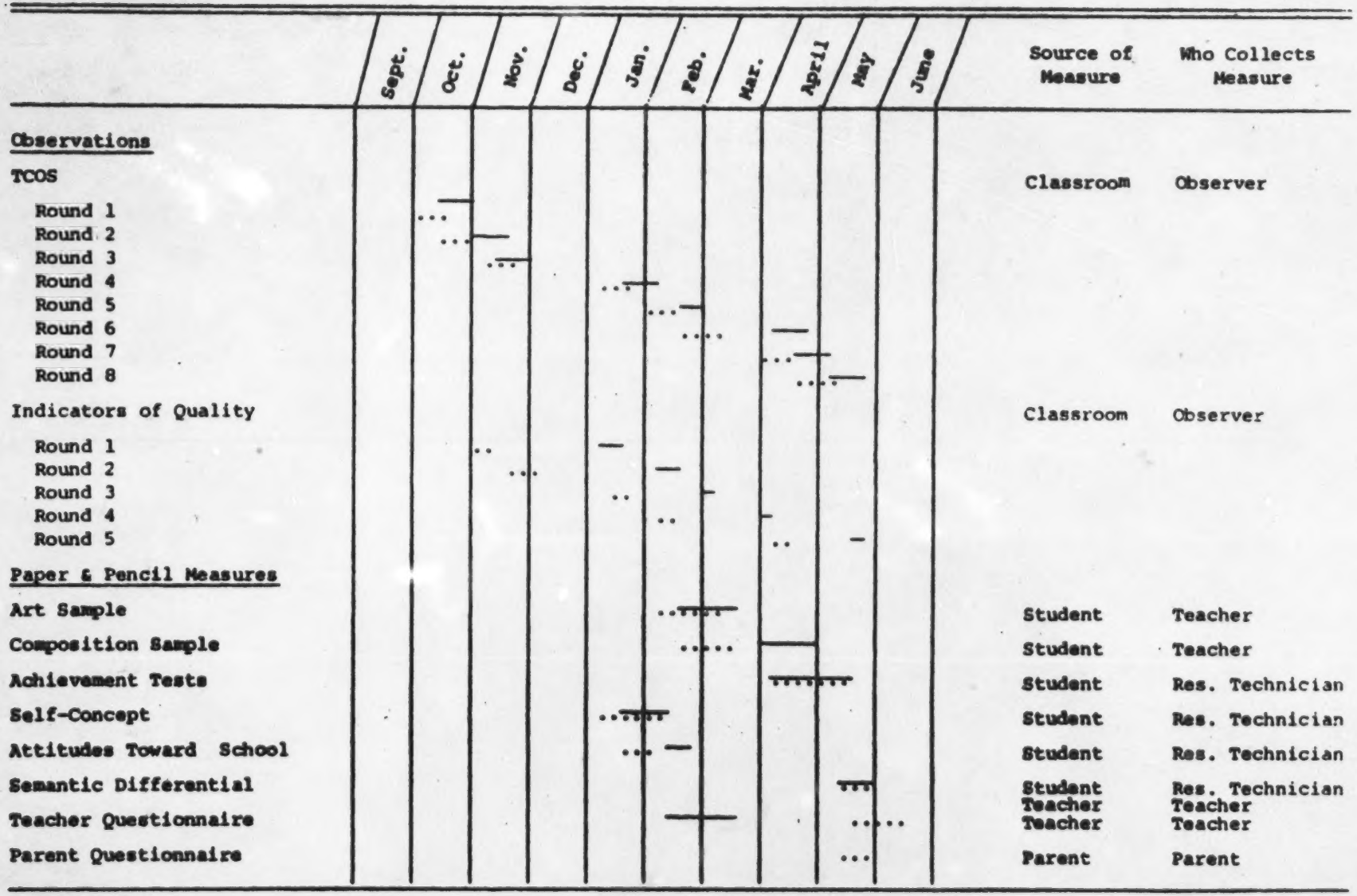
	Class Size			
	16	23	30	37
(a) Student Outcomes				
- academic improvement	12	6	0	0
- non-cognitive effects: more involvement, self-confidence	6	6	0	0
(b) Process/Environment and Teacher Effects				
- more teacher-pupil interaction, more individual attention to pupils' problems, better rapport	40	30	0	0
- more pupil-pupil interaction and group solidarity	9	7	1	0
- less formal teaching	2	1	0	0
- more relaxed environment, more materials and space	8	14	2	0
- better discipline	4	2	0	0
- time to make aids, change methods	2	0	0	0
(c) Program Effects				
- more individualized program	15	12	0	0
- more remedial help, enrichment	8	17	0	0
- more extracurricular activities, field trips and experience learning	10	7	0	0
- more small group work possible	3	7	1	0
(d) Other Comments	1	4	0	4
TOTAL	120	113	13	6

	Class Size			
	16	23	30	37
(a) Student Outcomes				
- decreases academically	0	0	3	3
- non-cognitive effects: little involvement, frustration, dependance	1	0	9	9
(b) Process-Environment and Teacher Effects				
- less teacher pupil contact, less individual attention, little time for pupils' problems	0	0	21	28
- class split into social groups; poor socially	1	0	1	1
- more structure, formality and seat work	0	0	8	11
- crowded environment, less space and equipment	0	0	1	3
- discipline difficult, less co-operation from pupils	1	2	1	4
- heavy workload, strenuous work situation	0	0	1	6
(c) Program Effects				
- more group-oriented program	0	0	3	8
- less remedial help, little in depth study	0	0	3	11
- limited extracurricular activities, field trips and experience learning				
- less group work possible or larger group required	1	0	7	8
(d) Other Comments	2	8	1	2
TOTAL	6	10	60	97

* Responses of 34 teachers to four open-ended questions regarding possible effects of each class size. Categories were combined so that the number of responses in a category may exceed the number of teachers.

APPENDIX B

TIMETABLE OF DATA COLLECTION ACTIVITIES IN THE SCHOOLS
(SEPTEMBER, 1974 - JUNE, 1976)



——— First year of the study (Grade 4, 1974-75).

..... Second year of the study (Grade 5, 1975-76).